UDK 811.112.2'374
Izvorni znanstveni rad
Rukopis primljen 30. IX. 2019.
Prihvaćen za tisak 23. I. 2020.
doi.org/10.31724/rihjj.46.2.25

Christine Möhrs
Sarah Torres Cajo
Leibniz-Institut für Deutsche Sprache (The Leibniz Institute for the German Language)
R 5, 6-13, DE-68161 Mannheim
moehrs@ids-mannheim.de
torres@ids-mannheim.de

THE MICROSTRUCTURE OF A LEXICOGRAPHICAL RESOURCE OF SPOKEN GERMAN: MEANINGS AND FUNCTIONS OF THE LEMMA *EBEN*[1]

This paper presents the corpus-based lexicographical prototype that was developed within the framework of the project *Lexik des gesprochenen Deutsch* (=LeGeDe) as a third-party funded project. Research results regarding the information offered in dictionaries have shown that there is a necessity for information on spoken lexis and its interactional functions. The resulting LeGeDe-prototype is based on these needs and desiderata and is thus an innovative example for the adequate representation of spoken language in online dictionaries. It is available online since September 2019 (https://www.owid.de/legede/). In the following sections, after first focusing on the presentation of the project's goals, the data basis, the intended end user, and the applied methods, we will illustrate the microstructure of the prototype and the information provided in a dictionary entry based on the lemma *eben*. Finally, we will summarize innovative aspects that are important for the implementation of such a resource.

1. Introduction

The resource this paper deals with was created within the framework of the third-party funded research project *Lexik des gesprochenen Deutsch* (LeGeDe[2])

[1] We thank Katja Arens, Luisa Cimander, Julia Hofmann and two anonymous reviewers for valuable commentaries on earlier drafts of this paper.

[2] www1.ids-mannheim.de/lexik/lexik-des-gesprochenen-deutsch.html [19 February 2020].

(eng. *Lexis of spoken German*) financed by the Leibniz Association (Leibniz Competition 2016, funding line 1: innovative projects[3]). The duration was three years and the project was carried out at the Leibniz-Institut für Deutsche Sprache (IDS) (eng. The Leibniz Institute for the German Language) in Mannheim as a cooperative project of the Department of Lexical Studies and the Department of Pragmatics (cf. Meliss and Möhrs 2017; Möhrs, Meliss and Batinić 2017). The collaboration of the two departments enabled the merging of the respective professional competences that are necessary for the development of a corpus-based lexicographical resource of spoken German in interaction (cf. the LeGeDe-prototype: https://www.owid.de/legede/). Thus, different disciplines could be represented in the project group: The team consisted of researchers with a special focus on lexicography, corpus linguistics, and interactional linguistics.

In this paper, we first give a brief overview of some basic assumptions and objectives that served as starting points for the project and pointed to the necessity of a lexicographical resource that deals with the peculiarities of the spoken language (see section 2). We briefly explain the database (the FOLK corpus) for the project analyses in section 3.1. In sections 3.2 and 3.3 we describe the corpus-based procedure for identifying the headword candidates and the analysis methods for developing the dictionary contents. Here, the lemma *eben* serves as an example to show how the dictionary entries are designed. For this part, we focus on one specific interactional function of *eben* (see section 3.3). Finally, the paper concludes by summarizing the advantages of the applied methods, the data used, and the outcome of the prototype developed in the LeGeDe-project (see section 4).

2. Initial Considerations

Our working hypotheses can be summarized with four basic assumptions:

(i) Firstly, we assumed that there are distinctions in the lexicon of spoken vs. written German at different levels – this has been postulated in research literature for several years (cf. Schwitalla 2012; Deppermann, Proske and Zeschel 2017; Günthner 2017; Bergmann 2017; Helmer and Deppermann 2017).

[3] www.leibniz-gemeinschaft.de/forschung/leibniz-wettbewerb/gefoerderte-vorhaben.html [19 February 2020].

(ii) Secondly, there is insufficient lexicographical documentation of interaction-typical peculiarities of the spoken-language lexicon of German (cf. Meliss and Möhrs 2018; Meliss, Möhrs and Ribeiro Silveira 2018). In dictionaries published to date, the information is primarily based on written-language usage. Whether the usage (especially with regard to example sentences) is based on authentic language material or was constructed by the lexicographers usually remains elusive (cf. Möhrs 2016: 351ff.). If these dictionaries contain information about usage in spoken-language contexts, this information also remains rather vague. Table 1 shows an excerpt from the dictionary entry of the lemma *eben* in the *Langenscheidt Großwörterbuch Deutsch als Fremdsprache* (eng. *Langenscheidt Dictionary for German as a Foreign Language*) in order to illustrate this assessment with an example. For some uses, only very vague references to prosodic characteristics are given (‘betont’/‘unbetont’, eng. ‘with/without emphasis’) or the label ‘gesprochen’ (eng. ‘spoken’) is assigned (these aspects are highlighted in bold, see Table 1).

Table 1: Lemma *eben* from the *Langenscheidt Dictionary for German as a Foreign Language*[4]

Langenscheidt Großwörterbuch Deutsch als Fremdsprache
1 ohne Berge oder Hügel und Täler <Land, eine Straße> ≈ flach **2** an allen Stellen gleichmäßig hoch <eine Fläche>
3 sehr kurz vor dem jetzigen Zeitpunkt **4** in diesem Augenblick ≈ jetzt **5** *eben (noch)* gerade noch zum richtigen Zeitpunkt **6** *norddeutsch,* ***gesprochen*** für (relativ) kurze Zeit ≈ schnell *Komm doch eben mal zu mir! (Expl. A)*

[4] We made the same observation in other printed and online dictionaries.

7 ***unbetont*** verwendet, um zu sagen, dass etwas nicht geändert werden kann und akzeptiert werden muss

Das ist eben nicht mehr zu ändern; *Du musst dich eben damit abfinden, dass er dich nicht mag (Expl. B)*

8 ***unbetont*** verwendet, um einen Vorschlag zu machen, den man für die Lösung des Problems hält

9 ***unbetont*** verwendet, um ein Wort, einen Sachverhalt o. Ä. besonders zu betonen ≈ genau

10 ***betont*** verwendet, um Zustimmung auszudrücken und dass man selbst schon früher das Gleiche gesagt hat

„Dann müssen wir die Sitzung auf morgen verschieben." – „Eben!"; „Ich finde ihn sehr unzuverlässig." – „Eben!" (Expl. C)

11 ***betont*** verwendet, um auf ironische Weise Zustimmung (und zugleich Ungeduld) auszudrücken

12 *(oder) eben 'nicht* verwendet, um eine Verneinung zu verstärken

The examples, marked with A, B and C in Table 1, convey to the recipient that these examples are from spoken interaction. In example A, this is suggested by the label 'spoken', in example C by the quotation marks as a signal of literal speech. There is no information on the interactional contexts and thus the question of authenticity of the data remains unanswered; the examples seem to be constructed (example B). For the dictionary user, the peculiarities of talk in interaction are described only superficially. There are a number of other examples which (can) demonstrate the rather inadequate collection and preparation of information about the use of spoken language in traditional dictionaries (cf. Meliss and Möhrs 2018).

To sum up, we believed that there is room for improvement in various aspects, e.g. regarding information on 1. prosody, 2. sequentiality, 3. context, 4. meaning (in authentic conversation), 5. interactive function, and with respect to examples 6. providing them from authentic conversations.

(iii) Thirdly, previous research has shown that there is a need for teaching methodologies and materials informed by systematic analysis of spoken German – both for native and non-native speakers of German (cf. Handwerker, Bäuerle and Sieberg 2016; Meliss 2016).

(iv) Finally, there are very few projects dedicated to spoken language from a lexicographical perspective. There is a small project on spoken Danish (cf. Hansen and Hansen 2012) that focuses on interjections and there are theoretical considerations on spoken Slovenian (cf. Verdonik and Sepesy Maučec 2017). Thus, the development of a dictionary prototype for the lexicon of spoken German was an innovative project.

The main interest lay in the phenomena of spoken German which speakers use more frequently or differently than in written German, e.g. in terms of meaning or function in interaction. Hence, the headword candidates of the LeGeDe-prototype are close to the standard variety (in comparison to other varieties) and as frequency-supported comparative analyses have shown, occur frequently in spoken German (cf. Meliss et al. 2018). Thus, the LeGeDe-prototype provides information on meaning and function for almost 30 lemmas: *Ahnung*/*keine Ahnung*, *eben* (description of three different lemma types), *gucken*/*guck mal*/*mal gucken*, *gut* (two different lemma types), *halt*, *ja* (two different lemma types), *kriegen* (two different lemma types), *schauen*/*schau mal*/*mal schauen*, *schön* (three different lemma types), *wissen*/*ich weiß nicht*/*weißt du*/*ich weiß*/*wer weiß*/*was weiß ich*.

When setting the lemma form, we differentiate between one-word and multi-word lemmas. The LeGeDe-prototype, for example, describes *Ahnung* (eng. *idea*) as a one-word lemma with its basic meanings. In addition, a functional description of the multi-word lemma *keine Ahnung* (eng. *no idea*) can be found in the prototype (see section 3.2). The lemma *eben*, on which this paper focuses, is a one-word lemma, which is listed both with a description of meaning as an adverb and with a functional description as a modal and discourse particle (see section 3.3).

Since the LeGeDe-prototype is primarily intended to function as a knowledge store and vocabulary documentation, the lexicographic descriptions are initially aimed at a scientifically interested group of users (including conversation analysists, interaction linguists, corpus linguists, lexicologists, and lexicographers). However, the results of our empirical surveys on expectations of the resource have also shown that users in learning situations – especially in production situations – could benefit from the LeGeDe-prototype if experts (scientists, teachers, etc.) occupy an appropriate intermediary position (cf. Meliss, Möhrs and

Ribeiro Silveira 2018: 132). Based on the data provided, learning materials for the treatment of specific lexical phenomena of spoken interaction could be created for the use in lessons for German (as a foreign language).

In addition to our main goal – the development of a corpus-based lexicographical resource of the lexicon of spoken German, we formulated the following subgoals:

– The determination of the specific features of spoken language usage on different levels (form, content/function, conversational setting, etc.), with a focus on lexical specifics.

– The development of further (corpus-)linguistic methods for analyzing and structuring spoken language data.

– The development of innovative types of lexicographical information, which refer to the function of lexical units in interactional contexts.

– The integration of the lexicographical prototype into the dictionary portal OWID^{plus}[5].

For this paper, we selected one of our headword candidates (*eben*), on which we will focus more closely. This example is intended to illustrate some of the goals just mentioned.

3. Insights into the Project Work

3.1. Data Basis

The data basis for the LeGeDe-project is FOLK[6] (*Forschungs- und Lehrkorpus Gesprochenes Deutsch*, eng. *Research and Teaching Corpus of Spoken German*). FOLK is part of the *Datenbank für gesprochenes Deutsch* (=DGD[7]) (eng. *Database for Spoken German*) and contains conversations and transcripts from

[5] "OWID^{plus} is a new experimental platform for multilingual lexical-lexicographic data, for quantitative lexical analyses and for interactive lexical applications which are presented in individual, unconnected sections. OWID^{plus} addresses an academic audience and is under active development."; cf. URL: www.owid.de/plus [19 February 2020].

[6] http://agd.ids-mannheim.de/folk.shtml [19 February 2020].

[7] https://dgd.ids-mannheim.de [19 February 2020].

German-speaking countries in various private, institutional, and public contexts, with currently about 250 hours of recorded material or about 2,4 mio. tokens respectively (release 2.12, [May 2019], cf. Schmidt 2014a, b). For selected aspects, we used comparative data on written German from DeReKo (*Das Deutsche Referenzkorpus*, eng. *The Mannheim German Reference Corpus*, cf. Kupietz and Keibel 2009).

3.2. Methodological Background

As a first quantitative step, we identified potential headword candidates with the help of frequency class comparisons from FOLK and a subset from DeReKo. The extraction was carried out for one-word lemmas. Manual qualitative analyses of a specifically defined sample from FOLK finally led to multi-word headword candidates, which were reflected on the one hand on the basis of results from research literature and on the other hand with results from quantitative calculations on bi- and trigrams. These candidates were particularly characterized by the fact that they have special interactional function and can be described as a whole unit (e.g. *weißt du*, eng. *you know*). With this combined procedure, we selected suitable and relevant headword candidates for the LeGeDe-prototype and described them in detail: For both one-word and multi-word lemmas, the results of meaning-related and function-related analyses were incorporated into the dictionary entries.

For each one-word lemma, we drew a random sample of 300 occurrences of which we analyzed 100 valid cases, employing a mixed-methods-approach of qualitative and quantitative analyses (cf. Johnson, Onwuegbuzie and Turner 2007). We used both meaning-based lexicological analyses, such as those tested in corpus-based dictionary projects (cf. Klosa 2011), as well as interactional analyses (cf. Sacks, Schegloff and Jefferson 1974; Selting and Couper-Kuhlen 2018) to disambiguate the meanings and functions of a lemma. In the LeGeDe-prototype, we provide three different types of articles: (i) Since many lemmas naturally show classical lexicological characteristics and semantic meaning, they are described in a meaning-based type of article. In this more traditional article, we provide information on the basic meanings of a lemma using authentic interactional data (transcripts and audio) from FOLK. (ii) Lemmas which primarily show interac-

tional functions are described in a function-based type of article. Here we provide information on the interactional functions of the lemma in context, also by applying authentic examples as transcripts with linked audio-files. Additionally, contextual, sequential, syntactic, and prosodic attributes are illustrated here. (iii) For a better overview of the rather complex meanings and functions as well as their interrelation, we finally provide a summary article, which serves as a link between the meaning- and function-based articles. Here the user can read about the general use of a lemma in spoken German and is able to navigate to different articles, semantic meanings, and interactional functions.

In order to explain the microstructure of the LeGeDe-prototype, we now focus on the function-based type of article. In our sample, the lexical item *eben* appeared in four word categories: modal particle (n=67), adverb (n=20), discourse particle (n=12), and adjective (n=1). For the article about *eben* as a modal particle (=MP), we coded all relevant cases (i.e. 67 MP-occurrences)[8], employing a coding scheme especially developed by using qualitative single-case analyses to work out relevant coding parameters. These parameters include communicative actions, pre- and post-context, sequential parameters, co-occurrences, prosodic realizations, and interactive functions.[9] On the basis of qualitative as well as quantitative data, we were able to work out the specific meanings and functions of a lemma and draw quantifiable conclusions.

3.3. Microstructure of the LeGeDe-Prototype: The Lemma *eben*

For *eben*, we were able to work out two semantic meanings and three interactional functions: (i) As an adverb, *eben* can be used in a sense of 'earlier' or 'with little effort', referring to either a temporal or action-based level. (ii) As a modal particle, *eben* can help to achieve coherence or plausibility, in both cases on a very argumentative level, whereas (iii) *eben* as a discourse particle is used to reclaim epistemic priority (cf. Betz and Deppermann 2018).

[8] For this paper, we only present the interactional functions of *eben* as a modal particle. For an overview of *eben* in other word categories see the LeGeDe-prototype as well as specifically Betz and Deppermann 2018 for an overview of *eben* as a discourse particle. Within the LeGeDe-prototype, Christine Möhrs worked on the parts in the dictionary article regarding the meaning (adverb) of *eben*, while Sarah Torres Cajo described the functions of *eben* as a modal particle and as a discourse particle.

[9] For the meaning-based type of article we also use a coding scheme with similar but slightly different parameters, e.g. semantic meanings and grammatical information. For a more detailed overview see Meliss, Möhrs, Ribeiro Silveira and Schmidt 2019.

Figure 1 shows an excerpt from the dictionary entry *eben*, in which the basic structure of an article entry becomes clear.

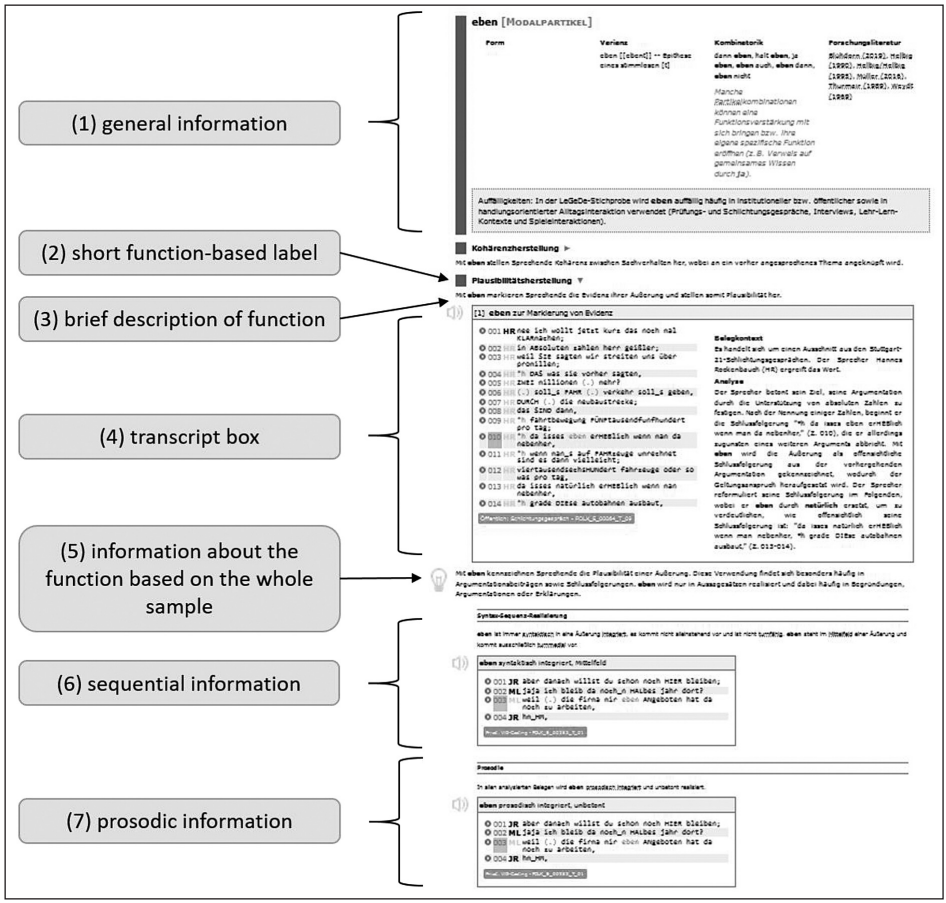

Figure 1: Excerpt from the dictionary entry *eben* as an MP

At the very top of every article entry we provide general information on the lemma (1), such as formal information, data on variation and combinatorics as well as relevant scientific research. Below that, each function is listed individually, identified by a short function-based label (2), and a brief description of the respective function (3).

We now focus on the interactional function of achieving coherence for which *eben* is recurrently used in social interaction. Modal particles function as con-

nective elements and "there is general agreement on the relational, respectively indexical nature of MPs" (Diewald 2006: 415f.). Some studies already ascribed a specific connective function to *eben* by arguing that it builds coherence by referring to a pragmatic pretext (cf. Thurmair 1989; Blühdorn 2019). Our data shows that, while this is true, *eben* goes beyond that as it can e.g. be used to establish an argumentative thread and is therefore a rhetorical as well as a turn-building element. Interlocutors use *eben* recurrently to return to their main sequence after an insertion (this equals (3) in Figure 1). In our prototype, we illustrate this by means of an exemplary case analysis, which we present in a transcript box (4), consisting of a transcript excerpt[10] (a) including a link to FOLK (b), context information (c), and the case analysis (d) (cf. Figure 2).

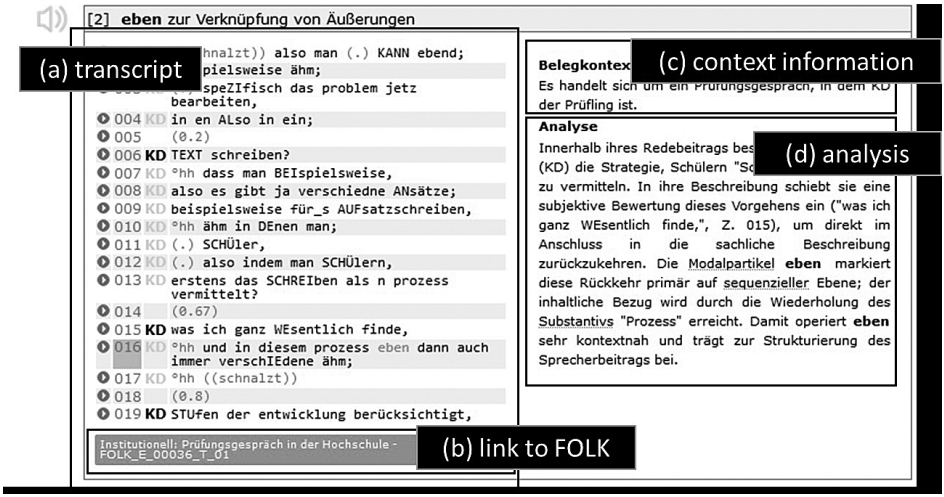

Figure 2: Transcript box

In the following extract (cf. Example 1), a museum guide (=BB) describes the structure and characteristics of a petrified skeleton of a Tyrannosaurus Rex (named "Tristan Otto"). The group has just arrived at the new exhibition object.

[10] In order to get access to the aligned audio, it is necessary to register with the DGD.

Example 1: Building coherence
[FOLK_E_00313_SE_01_T_02_DF_01_c323][11]

001 **BB** jetz sieht dieses gestein ganz ANders aus
now this rock looks completely different
002 BB als das geSTEIN
than the rock
003 BB was wir VORne haben
we have up front
004 BB DAS hier
this one
005 BB das sieht sehr DUNkel aus
this looks very dark
006 BB und äh tristan OTto wie der jetz hier geNANNT wurde von den besItzern
and uh tristan otto how he has been named by the owners
007 BB der WURde
he was
008 BB in einer geSTEINSformation gefunden
found in a geological formation
009 BB die nennt sich hell creek forMAtion
which is called hell creek formation
010 BB er wurde in monTAna gefunden
he was found in montana
011 BB und das is **eben** ne andre geSTEINSart
*and this is **EBEN** a different type of rock*
012 BB das is ein TONgestein
it's clay rock

The guide introduces the skeleton in contrast to a previously discussed type of rock by comparing them to each other. He begins a detailed description of the rock and focuses on the appearance, especially the color: "this one this looks very dark" (l. 004f.). With the connector 'and', the speaker initiates a parenthetical sequence in which he gives background information on the skeleton specifi-

[11] All transcripts follow the transcription-guidelines of GAT2 outlined in Selting et al. 2009. Capitalized syllables mark the main accentuation as indicated by prosodic accent or volume. The non-literal English translation is provided for each line individually and the target lemma is highlighted in bold letters and capitalized. The transcripts provided in the dictionary show more prosodic information, such as intonation and pitch. For the purpose of this paper, we simplified the notation to increase intelligibility. For a more detailed analysis of this extract in the context of an exhaustive description of *eben* as an MP, see Torres Cajo 2019.

cally, i.e. its name and the archaeological site it was found in. This digression from the primary topic (the rock in general) to the skeleton specifically is first marked by explicitly changing the subject to "tristan otto how he was named by the owners" (l. 006). Resulting from this, there is a change of pronouns (*das* to *der/er*, l. 006), which further indicates the change of focus (*Fokuswechsel*, cf. Kallmeyer 1987). This acts as a means to provide additional information on the exhibition object without leaving the knowledge-mediating sequence. The return to the main sequence is then indicated first by changing the pronoun back to "this" (l. 011) and second by the insertion of the MP *eben*: "and this is EBEN a different type of rock" (l. 011). This links the utterance with the previously begun description of the rock as darker than the one mentioned before. Thus, *eben* retrospectively establishes a coherent connection between the two segments, so the speaker can return to the main sequence and continue his explanation, finally solving the mystery and telling the audience that this is a "clay rock" (l. 012). The seemingly abrupt change of pronoun and focus back to the main sequence are comprehensible and understandable for the visitors as nobody shows any signs of misunderstanding or confusion. Therefore, the modal particle *eben* contributes to structuring and organizing the guide's longer contribution in such a way that he can smoothly provide explanation on various levels without becoming incoherent or risking misunderstanding on the part of the audience. Additionally, the line of argument is effortlessly and coherently kept together by the turn-constructional components.

In the dictionary entry, this detailed case analysis is followed by general information on the respective function, which refers to the results of the entire coded sample (=(5) in Figure 1). Throughout the whole sample, we can see that *eben* is recurrently used to retrospectively connect utterances on a content or sequential level and hence establishes coherence within a speaker's turn(s). This can be done either over longer stretches of talk (one extract which is used for illustration in the dictionary shows a side sequence of more than six minutes) or in close sequential context (as we have seen in Example 1). Usually, interlocutors use *eben* to refer to a fact or an utterance previously expressed by themselves. In rare cases, it is also possible to refer to other participants' earlier statements. We also noticed the frequent co-occurrence *eben nicht* (*eben not*), which acts as a means to emphasize an argumentative context. We can see in our data that *eben* is systematically used by participants as a strategic element to increase the sustainability of a state-

ment and build a coherent argumentative thread, especially in such interactions in which arguing is a central communicative task (e.g. oral exams, mediation talks, educational settings).

The coding of 100 appearances of a lemma enabled us to make quantitative as well as overarching statements regarding syntactic (6), sequential (6), and prosodic parameters (7). In regard to sequential realizations, we can see that *eben* appears as a syntactically integrated element, mostly in the middle field[12] of a sentence (cf. Example 2). With regard to the description of syntactic, sequential, and prosodic features, we also provide a transcript box, but without a detailed context- and functional analysis as we only want to illustrate possible sequential realization patterns.

Example 2: Syntactic information
[FOLK_E_00015_SE_01_T_01_DF_01_c405]

001	CH	WENN sie *when you*
002		(0.29)
003	CH	äh DAten (.) analysieren *uh analyze data*
004	CH	(.) isses HÄUfig **eben** schwierig zu sagen *it is **EBEN** often difficult to say*

In spoken interaction, some utterances do not form a full sentence, but constitute expansions or fragments (cf. Auer 2000 for the concept of *on-line syntax*). In these more infrequent utterances (cf. Example 3), *eben* appears in the periphery (cf. Imo 2008, who describes this phenomenon for the MP *halt*).

[12] This terminology refers to the concept of the topological field model (*topologisches Feldermodell*), which grasps German syntactic structures.

Example 3: Syntactic information 2
[FOLK_E_00042_SE_01_T_02_DF_01_c897]

001 LS aber des stimmt SCHON _ n bisschen mit der dankbarkeit dass sie dann das gefühl haben sie müssen was zurückgeben auf ihre art und weise **eben**

*but it's true in a way concerning the gratitude, that they think they have to give something back in their own way **EBEN***

Concerning prosody, we can see that *eben* always occurs without a prosodic accent and is always prosodically integrated into the intonation contour. This confirms the general assumption that modal particles usually do not have a prosodic accent (cf. Thurmair 1989). As we cannot show any audio data here, we are not able to give a sufficient example for the prosodic realization. In our dictionary, we provide transcripts with embedded audio-files which illustrate the given statements (see (7) in Figure 1).

In addition to results about interactional function, syntax, and prosody, we can lay out further information regarding metadata, such as interactional settings or participation framework. For *eben,* we can see that it often occurs in the context of institutional or public interaction as well as action-based everyday-interaction, e.g. panel discussions, conciliation talks, team meetings or private game-playing interactions. This seems to be related to the argumentative and coherence-building effects of the MP: As mentioned before, *eben* helps the speaker to build argumentative threads and acts as a turn-structuring element, which supports the organization of longer sequences in these interactions.

4. Concluding Remarks

Since we were the first to develop a corpus-based lexicographical resource on spoken German, a number of methodological and technical challenges were involved during the course of our project. These included the quantitative identification of relevant headword candidates using a specially developed methodological approach, qualitative analysis and the subsequent structuring of the data for lexicographical implementation. The spectrum of lexicographical content,

which combines classical data with a novel information offer, is multimodal and was developed specifically for the description of lexical phenomena in interaction. Using the example of the lemma *eben*, we were able to demonstrate the potential of the lexicographic description illustrated in the LeGeDe-prototype. Additionally, we integrated authentic corpus data into the dictionary articles via selected transcript excerpts and aligned audio data. The LeGeDe-prototype is one of the few lexicographical resources that have a direct link to the corresponding corpus (FOLK). This allows the user to get access to the whole conversation, to learn more about the metadata of the interaction, and to research further into the FOLK data.

Since the development of a corpus-based dictionary on spoken German is a novum that was achieved with adequate methods to analyze, structure, and describe items of verbal speech, it can serve as an example and conduce to similar research projects for other languages. Thus, the project fosters the scientific exchange in the field of linguistics and represents a new milestone in the lexicographic area.

References

Auer, Peter. 2000. Online-Syntax – Oder: was es bedeuten könnte, die Zeitlichkeit der mündlichen Sprache ernst zu nehmen. *Sprache und Literatur* 85. 43–56. doi-org.ubproxy.ub.uni-heidelberg.de/10.30965/25890859-031-01-90000005.

Bergmann, Pia. 2017. Gebrauchsprofile von *weiß nich* und *keine Ahnung* im Gespräch. Ein Blick auf nicht-responsive Vorkommen. *Diskursmarker im Deutschen. Reflexionen und Analysen*. Eds. Blühdorn, Hardarik; Deppermann, Arnulf; Helmer, Henrike; Spranz-Fogasy, Thomas. Verlag für Gesprächsforschung. Göttingen. 157–182. http://d-nb.info/1147303339 (accessed 25 February 2020).

Betz, Emma; Deppermann, Arnulf. 2018. Indexing Priority of Position: *Eben* as Response Particle in German. *Research on Language and Social Interaction* 51/2. 171–193. doi.org/10.1080/08351813.2018.1449449.

Blühdorn, Hardarik. 2019. Modalpartikeln und Akzent im Deutschen. *Linguistische Berichte* 259. 275–317.

Diewald, Gabriele. 2006. Discourse particles and modal particles as grammatical elements. *Approaches to discourse particles*. Ed. Fischer, Kerstin. Elsevier. Amsterdam. 403–426.

elexiko. Erfahrungsberichte aus der lexikografischen Praxis eines Internetwörterbuchs. 2011. Ed. Klosa, Annette. Narr. Tübingen. (*Studien zur Deutschen Sprache* 55). nbn-resolving.org/urn:nbn:de:bsz:mh39-51544 (accessed 25 February 2020).

Gesprochene Fremdsprache Deutsch. 2016. Eds. Handwerker, Brigitte; Bäuerle, Rainer; Sieberg, Bernd. Schneider. Baltmannsweiler. (*Perspektiven Deutsch als Fremdsprache* 32). http://d-nb.info/1093201932 (accessed 25 February 2020).

Günthner, Susanne. 2017. Diskursmarker in der Interaktion – Formen und Funktionen univerbierter *guck mal*- und *weißt du*-Konstruktionen. *Diskursmarker im Deutschen. Reflexionen und Analysen*. Eds. Blühdorn, Hardarik et al. Verlag für Gesprächsforschung. Göttingen. 103–130. http://d-nb.info/1147303339 (accessed 25 February 2020).

FOLK – Forschungs- und Lehrkorpus Gesprochenes Deutsch. 2020. Leibniz-Institut für Deutsche Sprache. Mannheim. http://agd.ids-mannheim.de/folk.shtml (accessed 19 February 2020).

Helmer, Henrike; Deppermann, Arnulf. 2017. ICH WEIß NICHT zwischen Assertion und Diskursmarker: Verwendungsspektren eines Ausdrucks und Überlegungen zu Kriterien für Diskursmarker. *Diskursmarker im Deutschen Reflexionen und Analysen*. Eds. Blühdorn, Hardarik et al. Verlag für Gesprächsforschung. Göttingen. 131–156. http://d-nb.info/1147303339 (accessed 25 February 2020).

Imo, Wolfgang. 2008. Individuelle Konstrukte oder Vorboten einer neuen Konstruktion? Stellungsvarianten der Modalpartikel *halt* im Vor- und Nachfeld. *Konstruktionsgrammatik II: Von der Konstruktion zur Grammatik.* Eds. Stefanowitsch, Anatol; Fischer, Kerstin. Stauffenburg. Tübingen. 135–156. http://d-nb.info/988811545 (accessed 25 February 2020).

Johnson, Burke; Onwuegbuzie, Anthony; Turner, Lisa. 2007. Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research* 1/2. 112–133. doi.org/10.1177/1558689806298224.

Kallmeyer, Werner. 1987. Bemerkungen zum Verhältnis von Sprache und Interaktion am Beispiel von ‚Ja aber'. Sprache und Gesellschaft in deutschsprachigen Ländern. *Beiträge eines Werkstattgesprächs des Goethe-Instituts York in Zusammenarbeit mit dem Department of Language and Linguistic Science der Universität York am 16./17.9.85 in York.* Eds. Russ, Charles; Volkmar, Claudia. Goethe-Institut. München. 53–74.

Kupietz, Marc; Keibel, Holger. 2009. The Mannheim German Reference Corpus (DeReKo) as a basis for empirical linguistic research. *Working Papers in Corpus-based Linguistics and Language Education.* Eds. Minegishi, Makoto; Kawaguchi, Yuji. Tokyo University of Foreign Studies (=TUFS). Tokyo. 53–59. http://cblle.tufs.ac.jp/assets/files/publications/working_papers_03/section/053-059.pdf (accessed 19 February 2020).

Leibniz Gemeinschaft. 2020. Im Leibniz-Wettbewerb geförderte Vorhaben. www.leibniz-gemeinschaft.de/forschung/leibniz-wettbewerb/gefoerderte-vorhaben.html (accessed 19 February 2020).

LeGeDe-prototype. www.owid.de/legede (accessed 24 September 2019).

LeGeDe – Lexik des gesprochenen Deutsch. Project Website. 2020. Leibniz-Institut für Deutsche Sprache. Mannheim. http://www1.ids-mannheim.de/lexik/lexik-des-gesprochenen-deutsch.html (accessed 19 February 2020).

Meliss, Meike. 2016. Gesprochene Sprache in DaF-Lernerwörterbüchern. *Gesprochene Fremdsprache Deutsch*. Eds. Handwerker, Brigitte; Bäuerle, Rainer; Sieberg, Bernd. Schneider. Baltmannsweiler. 179–199. http://d-nb.info/1093201932 (accessed 25 February 2020).

Meliss, Meike; Möhrs, Christine. 2017. Die Entwicklung einer lexikografischen Ressource im Rahmen des Projektes LeGeDe. *Sprachreport. Informationen und Meinungen zur deutschen Sprache* 4/17. 42–52. https://nbn-resolving.org/urn:nbn:de:bsz:mh39-68549 (accessed 25 February 2020).

Meliss, Meike; Möhrs, Christine. 2018. Lexik in der spontanen, gesprochensprachlichen Interaktion: Eine anwendungsorientierte Annäherung aus der DaF-Perspektive. *GFL* 3/2018. 79–110. www.gfl-journal.de/Issue_3_2018.php (accessed 19 February 2020.).

Meliss, Meike; Möhrs, Christine; Batinić, Dolores; Perkuhn, Rainer. 2018. Creating a list of headwords for a lexical resource of spoken German. *Proceedings of the 18th EURALEX International Congress 2018.* Eds. Čibej, Jaka; Gorjanc, Vojko; Kosem, Iztok; Krek, Simon. Znanstvena založba Filozofske fakultete Univerze v Ljubljani. Ljubljana. 1009–1016. https://nbn-resolving.org/urn:nbn:de:bsz:mh39-77208 (accessed 25 February 2020).

Meliss, Meike; Möhrs, Christine; Ribeiro Silveira, Maria. 2018. Erwartungen an eine korpusbasierte lexikografische Ressource zur „Lexik des gesprochenen Deutsch in der Interaktion“: Ergebnisse aus zwei empirischen Studien. *Zeitschrift für Angewandte Linguistik* 68. 103–138. doi.org/10.1515/zfal-2018-0005.

Meliss, Meike; Möhrs, Christine; Ribeiro Silveira, Maria; Schmidt, Thomas. 2019. A corpus-based lexical resource of spoken German in interaction. *Electronic Lexicography in the 21st century: Proceedings of the eLex-Conference 2019*. Eds. Kosem, Iztok et al. Lexical Computing CZ, s.r.o. Brno. 783–804. https://nbn-resolving.org/urn:nbn:de:bsz:mh39-93621 (accessed 25 February 2020).

Möhrs, christine. 2016. *Syntagmatische Verwendungsmuster in einsprachigen deutschen Wörterbüchern*. Institut für Deutsche Sprache. Mannheim. https://nbn-resolving.org/urn:nbn:de:bsz:mh39-58541 (accessed 25 February 2020).

Möhrs, Christine; Meliss, Meike; Batinić, Dolores. 2017. LeGeDe – towards a corpus-based lexical resource of spoken German. *Electronic Lexicography in the 21st century: Proceedings of eLex 2017 Conference*. Eds. Kosem, Iztok; Tiberius, Carole; Jakubíček, Miloš; Kallas, Jelena; Krek, Simon; Baisa, Vít. Lexical Computing CZ s.r.o. Brno.

281–298. https://nbn-resolving.org/urn:nbn:de:bsz:mh39-64214 (accessed 25 February 2020).

OWID^{plus} – Plattform für multilinguale lexikalisch-lexikografische Daten. Leibniz-Institut für Deutsche Sprache. Mannheim. www.owid.de/plus (accessed 19 February 2020).

Sacks, Harvey; Schegloff, Emanuel; Jefferson, Gail. 1974. A simplest systematics for the organization of turn-taking in conversation. *Language* 50/4–1. 696–735. doi.org/10.2307/412243.

Schmidt, Thomas. 2014a. The Research and Teaching Corpus of Spoken German – FOLK. *Proceedings of the 9th International Conference on Language Resources and Evaluation.* Eds. Calzolari, Nicoletta et al. European Language Resources Association. Reykjavik. 383–387. https://nbn-resolving.org/urn:nbn:de:bsz:mh39-24434 (accessed 25 February 2020).

Schmidt, Thomas. 2014b. Gesprächskorpora und Gesprächsdatenbanken am Beispiel von FOLK und DGD. *Gesprächsforschung. Online-Zeitschrift zur verbalen Interaktion* 15. 196–233. www.gespraechsforschung-online.de/2014.html (accessed 19 February 2020).

Schwitalla, Johannes. 2012. *Gesprochenes Deutsch. Eine Einführung.* Schmidt. Berlin. (*Grundlagen der Germanistik* 33). http://d-nb.info/1013377036 (accessed 25 February 2020).

Selting, Margret; Auer, Peter; Barth-Weingarten, Dagmar; Bergmann, Jörg; Bergmann, Pia; Birkner, Karin; Couper-Kuhlen, Elizabeth; Deppermann, Arnulf; Gilles, Peter; Günthner, Susanne; Hartung, Martin; Kern, Friederike; Mertzlufft, Christine; Meyer, Christian; Morek, Miriam; Oberzaucher, Frank; Peters, Jörg; Quasthoff, Uta; Schütte, Wilfried; Stukenbrock, Anja; Uhmann, Susanne. 2009. Gesprächsanalytisches Transkriptionssystem 2 (GAT2). *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 10. 353–402. www.gespraechsforschung-online.de/heft2009/heft2009.html (accessed 19 February 2020).

Selting, Margret; Couper-Kuhlen, Elizabeth. 2018. *Interactional Linguistics: Studying language in social interaction.* Cambridge University Press. Cambridge.

Thurmair, Maria. 1989. *Modalpartikeln und ihre Kombinationen*. Niemeyer. Tübingen. http://d-nb.info/890200106 (accessed 25 February 2020).

Torres Cajo, Sarah. 2019. Zwischen Strukturierung, Wissensmanagement und Argumentation im Gespräch – Interaktionale Verwendungsweisen der Modalpartikeln *halt* und *eben* im gesprochenen Deutsch. *Deutsche Sprache* 4/47. 289–310. https://nbn-resolving.org/urn:nbn:de:bsz:mh39-93708 (accessed 25 February 2020).

Verben im interaktiven Kontext: Bewegungsverben und mentale Verben im gesprochenen Deutsch. 2017. Eds. Deppermann, Arnulf; Proske, Nadine; Zeschel, Arne.

Narr. Tübingen. https://nbn-resolving.org/urn:nbn:de:bsz:mh39-63314 (accessed 25 February 2020).

Verdonik, Darinka; Sepesy Maučec, Mirjam. 2017. A speech corpus as a source of lexical information. *International Journal of Lexicography* 30/2. 143–166. doi-org.ubproxy.ub.uni-heidelberg.de/10.1093/ijl/ecw004.

Mikrostruktura leksikografskih izvora govornoga njemačkog: značenja i funkcije natuknice *eben*

Sažetak

U radu se prikazuju struktura, funkcije i mogućnosti LeGeDe-prototipa analizirajući rječnički članak natuknice *eben*. Taj elektronički izvor pruža podatke o jedinicama govornoga njemačkog te daje pregled bogatstva značenja i interakcijskih funkcija različitih natuknica. Stoga je on vrijedan izvor u različitim znanstvenim područjima, primjerice u primijenjenoj lingvistici, te se može upotrebljavati pri oblikovanju obrazovnih materijala. Budući da je razvoj korpusno utemeljenoga rječnika govornoga njemačkog novost utemeljena na uporabi primjerenih metoda za analizu, strukturu i opis jedinica govornoga jezika, može se upotrijebiti kao primjer i putokaz za slične projekte drugih jezika. Stoga projekt njeguje razmjenu znanstvenih podataka u području jezikoslovlja te je prekretnica u leksikografiji.

Keywords: online dictionary, Interactional Linguistics, lexicographical resource, spoken German in interaction, microstructure, modal particles

Ključne riječi: mrežni rječnik, interakcijsko jezikoslovlje, leksikografski izvori, govorni njemački u interakciji, mikrostruktura, modalne čestice